CLAIMS

1. A method for identifying block/zone misalignments in a microarray data set, the method comprising:

calculating displacement vectors and displacement-vector-based metrics for the microarray data set comprising a digital image of a microarray and an initial feature indexing; and

partitioning the microarray data set into smaller partitions until each block/zone misalignment is isolated in one or more partitions, as determined by displacement-vector-based metrics calculated for each partition.

2. The method of claim 1 wherein displacement-vector-based metrics include:

a vector sum of the displacement vectors for a partition within a microarray calculated by summing all displacement vectors $\mathbf{d}_i$ in the partition according to

$$\boldsymbol{\mu}_v = \sum_{i=1}^{n} \mathbf{d}_i$$

where $\boldsymbol{\mu}_v$ is the a vector sum of the displacement vectors and n is the number of features;

a length of the vector sum of the displacement vectors within the partition calculated according to

$$\bar{\mu}_v = \sqrt{\boldsymbol{\mu}_v \cdot \boldsymbol{\mu}_v}$$

where $\bar{\mu}_v$ is the length of the vector sum of the displacement vectors; and

an average length of the displacement vectors $\mathbf{d}_i$ within the partition, calculated according to

$$\bar{\mu}_s = \frac{1}{n} \sum_{i=1}^{n} \sqrt{\mathbf{d}_i \cdot \mathbf{d}_i}$$

where $\bar{\mu}_s$ is the average length of the displacement vectors.

3. The method of claim 2 wherein partitioning the microarray data set into smaller partitions until block/zone misalignments are each isolated in one or more partitions further includes:

partitioning the microarray data into partitions until one of

$\bar{\mu}_s$ calculated for a partition falls below a threshold value,

$\bar{\mu}_s$ calculated for a final partition is less than or equal to $\bar{\mu}_s$ calculated for a larger partition that contains the final partition, or

$\frac{\bar{\mu}_v}{\bar{\mu}_s}$ calculated for a smaller partition is greater than $\frac{\bar{\mu}_v}{\bar{\mu}_s}$ calculated for a larger partition that contains the smaller partition.

4. The method of claim 2 further including:

following partitioning of the microarray data set, coalescing individual partitions that, when merged, have a combined $\bar{\mu}_s$ less than or equal to $\bar{\mu}_s$ for each individual partition and have a combined ratio $\frac{\bar{\mu}_v}{\bar{\mu}_s}$ less than or equal to $\frac{\bar{\mu}_v}{\bar{\mu}_s}$ for each individual partition.

5. The method of claim 2 further including recalculating feature positions taking into account detected block/zone misalignments.

6. The method of claim 2 wherein displacement vectors are calculated by matrix operations that determine a matrix

$$\mathbf{M}=\begin{pmatrix} m_{xx} & m_{xy} & O_x \\ m_{yx} & m_{yy} & O_y \end{pmatrix}$$

from a matrix containing the centroids of found features **C** and an inverse of an index matrix **I** as follows:

$\mathbf{M}=\mathbf{C}\mathbf{I}^{-1}$

where $\mathbf{C}=\begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \\ \cdots & \cdots \\ x_n & y_n \end{pmatrix}$ and $\mathbf{I}=\begin{pmatrix} 1\ 1\ 1\ \ldots\ 1 \\ c_1\ c_2\ c_3 \ldots\ c_n \\ r_1\ r_2\ r_3 \ldots\ r_n \end{pmatrix}$.

7. A method comprising forwarding to a remote location one of:

a block/zone misalignment determined by the method of claim 1;

feature position data obtained by correcting for a block/zone misalignment determined by the method of claim 1; and

results obtained using a microarray data set and feature position data obtained by correcting for a block/zone misalignment determined by the method of claim 1.

8. A computer program implementing the method of claim 1 stored in a computer-readable medium.

9. A microarray data processing system that performs the method of claim 1.

10. A microarray data processing system comprising:

stored microarray data set;

a processor; and

a program executed by the processor that

calculates displacement vectors and displacement-vector-based metrics for the microarray data set comprising a digital image of a microarray and an initial feature indexing; and

partitions the microarray data set into smaller partitions until block/zone misalignments are each isolated in one or more partitions, as determined by displacement-vector-based metrics calculated for each partition.

11. The microarray data processing system of claim 11 wherein displacement-vector-based metrics include:

a vector sum of the displacement vectors for a partition within a microarray calculated by summing all displacement vectors $\mathbf{d}_i$ in the partition according to

$$\boldsymbol{\mu}_v = \sum_{i=1}^{n} \mathbf{d}_i$$

where $\boldsymbol{\mu}_v$ is the a vector sum of the displacement vectors and n is the number of features;

a length of the vector sum of the displacement vectors within the partition calculated according to

$$\bar{\mu}_v = \sqrt{\boldsymbol{\mu}_v \cdot \boldsymbol{\mu}_v}$$

where $\bar{\mu}_v$ is the length of the vector sum of the displacement vectors; and

an average length of the displacement vectors $\mathbf{d}_i$ within the partition, calculated according to

$$\bar{\mu}_s = \frac{1}{n} \sum_{i=1}^{n} \sqrt{\mathbf{d}_i \cdot \mathbf{d}_i}$$

where $\bar{\mu}_s$ is the average length of the displacement vectors.

12. The microarray data processing system of claim 11 wherein the computer program partitions the microarray data set into smaller partitions until block/zone misalignments are each isolated in one or more partitions by:

partitioning the microarray data into partitions until one of

$\bar{\mu}_s$ calculated for a partition falls below a threshold value,

$\bar{\mu}_s$ calculated for a final partition is less than or equal to $\bar{\mu}_s$ calculated for a larger partition that contains the final partition; and

$\frac{\bar{\boldsymbol{\mu}}_v}{\bar{\mu}_s}$ calculated for a smaller partition is greater than $\frac{\bar{\boldsymbol{\mu}}_v}{\bar{\mu}_s}$ calculated for a larger partition that contains the smaller partition.

13. The microarray data processing system of claim 11 wherein the computer program, following partitioning of the microarray data set, coalesces individual partitions that, when

merged, have a combined $\bar{\mu}_s$ less than or equal to $\bar{\mu}_s$ for each individual partition and have a combined ratio $\frac{\bar{\mu}_v}{\bar{\mu}_s}$ less than or equal to $\frac{\bar{\mu}_v}{\bar{\mu}_s}$ for each individual partition.

14. A method for correcting block/zone misalignments in a microarray data set including indexed features, the method comprising:

calculating displacement vectors for the features; and

identifying block/zone misalignments within the features in the microarray data set using the calculated displacement vector.

15. The method of claim 14 wherein identifying block/zone misalignments within the features in the microarray data set further comprises:

partitioning the microarray data set into smaller partitions until block/zone misalignments are each isolated in one or more partitions, as determined by displacement-vector-based metrics calculated for each partition.

16. The method of claim 14 wherein displacement vectors are calculated by matrix operations that determine a matrix

$$\mathbf{M}=\begin{pmatrix} m_{xx} & m_{xy} & O_x \\ m_{yx} & m_{yy} & O_y \end{pmatrix}$$

from a matrix containing the centroids of found features **C** and an inverse of an index matrix **I** as follows:

$$\mathbf{M}=\mathbf{C}\mathbf{I}^{-1}$$

where $\mathbf{C}=\begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \\ \dots & \dots \\ x_n & y_n \end{pmatrix}$ and $\mathbf{I}=\begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ c_1 & c_2 & c_3 & \dots & c_n \\ r_1 & r_2 & r_3 & \dots & r_n \end{pmatrix}$.

17. The method of claim 16 further including generating corrected feature positions for the features of the block/zone misalignments by determining an estimate for the correct

positions of all features $\mathbf{F}_{all}$ within a partition, based on the matrix **M** determined using the observed positions of a set of known features, as follows:

$$\mathbf{F}_{all} = \mathbf{M}\mathbf{I}_{all}$$

18. A method comprising forwarding to a remote location one of:

feature position data obtained by correcting for a block or zone misalignment determined by the method of claim 14; and

results obtained using a microarray data set and feature position data obtained by correcting for a block or zone misalignment determined by the method of claim 14.

19. A computer program implementing the method of claim 14 stored in a computer-readable medium.

20. A microarray data processing system that performs the method of claim 14.